SEQUENCE LISTING

.5	(1) GENER	RAL INFORMATION:
	(i)	APPLICANT: Gorski, David H. Walsh, Kenneth
10	(ii)	TITLE OF INVENTION: Growth Arrest Homeobox Gene
10	(iii)	NUMBER OF SEQUENCES: 4
15	(iv)	CORRESPONDENCE ADDRESS: (A) ADDRESSEE: Calfee, Halter, and Griswold (B) STREET: 800 Superior Avenue (C) CITY: Cleveland (D) STATE: Ohio (E) COUNTRY: U.S.A. (F) ZIP: 44114-2688
2 5 0	(V)	COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
30	(vi)	CURRENT APPLICATION DATA: (A) APPLICATION NUMBER: (B) FILING DATE: (C) CLASSIFICATION:
35	(viii)	ATTORNEY/AGENT INFORMATION: (A) NAME: Golrick, Mary E. (B) REGISTRATION NUMBER: 34829 (C) REFERENCE/DOCKET NUMBER: 22311/00114
40	(ix)	TELECOMMUNICATION INFORMATION: (A) TELEPHONE: (216) 622-8200 (B) TELEFAX: (216) 241-0816 (C) TELEX: 980499
	(2) INFO	RMATION FOR SEQ ID NO:1:
45	(i)	SEQUENCE CHARACTERISTICS:
		(A) LENGTH: 2244 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: both
50		(D) TOPOLOGY: linear
//	(ii)	MOLECULE TYPE: CDNA
. /	(iii)	HYPOTHETICAL: NO
55	(iv)	ANTI-SENSE: NO
60	(ix)	FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 1971108
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:1:
65	GTCAAGTC 60	FTT TATACGTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT

and field as the first and analysis and the second of the second of the second second field as the of the field of the

		CTGC 120	TCAA	AC (CGCG	CGGC	T T	TAC	ATTAG	GAG	TGAG	STGG	GGGA	\GAG1	rcc :	[AGG]	ATTTCT	
5		AGTG 180	AAAA	GT C	ACAC	CGCT	T GO	etgga	CTTT	. GGG	ACCI	TCG	TGAA	GTCI	TTC :	IGCT!	IGGAAG	
		CTGA 229	GACT	TG (CATGO													
10						Me	t GI	u Hi	s Pr	o Le	u Pn 5	e GI	у Су:	s Le	u Ar 1	Ξ	r	
		CCC 277	CAC	GCC	ACA	GCG	CAA	GGC	TTG	CAC	CCC	TTC	TCG	CAG	TCT	TCT	CTG	
15	·	Pro	His	Ala	Thr 15	Ala	Gln	Gly	Leu	His 20	Pro	Phe	Ser (Gln	Ser 25	Ser	Leu	
		GCC 325	CTC	CAT	GGA	AGA	TCT	GAC	CAC	ATG	TCC	TAC	CCC	GAA	CTC	TCC	ACA	
20		Ala	Leu	His 30	Gly	Arg	Ser	Asp	His 35	Met	Ser	Tyr	Pro (Glu 40	Leu	Ser	Thr	
S.		TCT 373	TCC	TCG	TCT	TGC	ATA	ATC	GCG	GGA	TAC	ccc	AAT	GAG	GAG	GGC	ATG	
1 2 5		Ser	Ser 45	Ser	Ser	СЛа	Ile	Ile 50	Ala	Gly	Tyr	Pro	Asn 55	Glu	Glu [.]	Gly	Met	
All Williams		TTT 421	GCC	AGC	CAG	CAT	CAC	AGG	GGG	CAC	CAC	CAC	CAC	CAC	CAC	CAC	CAC	
30	•	Phe 60	Ala	Ser	Gln	His	His 65	Arg	Gly	His	His	His 70	His	His	His	His	His 75	
		CAT	CAC	CAC	CAC	CAG	CAG	CAG	CAG	CAC	CAG	GCT	CTG	CAA	AGC	AAC	TGG	
3 5		His	His	His	His	Gln 80	Gln	Gln	Gln	His.	Gln 85	Ala	Leu	Gln	Ser	Asn 90	Trp	
	٠	CAC 517	CTC	CCG	CAG	ATG	TCC	TCC	CCG	CCA	AGC	GCG	GCC	CGG	CAC	AGC	CTT	
<u>4</u> 0			Leu	Pro	Gln 95	Met	Ser	Ser	Pro	Pro 100	Ser	Ala [.]	Ala	Arg	His 105	Ser	Leu	
		TGC 565	CTG	CAG	CCT	GAT	TCC	GGA	GGG	CCC	CCG	GAG	CTG	GGG	AGC	AGC	CCT	
45		Cys	Leu	Gln 110	Pro	Asp	Ser		Gly 115	Pro	Pro	Glu		Gly 120	Ser	Ser	Pro	
		CCG 613	GTC	CTG	TGC	TCC	AAC	TCT	TCT	AGC	CTG	GGC	TCC	AGC	ACC	CCG	ACC	
50		Pro	Val 125	Leu	Cys	Ser	Asn	Ser 130	Ser	Ser	Leu	Gly	Ser 135	Ser	Thr	Pro	Thr	
		GGA 661	GCC	GCG	TĢC	GCA	CCA	AGG	GAT	TAT	GGC	CGT	CAA	GCG	CTG	TCA	CCC	
55			Ala	Ala	СЛа	Ala	Pro 145	Arg	yab	Tyr	Gly	Arg 150	Gln	Ala	Leu	Ser	Pro 155	
		GCA 709	GAA	GTG	GAG	AAG	AGA	AGT	GGC	AGC	AAA	AGA	. AAA	AGC	GAC	AGI	TCA	
60			Glu	Val	Glu	Lys 160	Arg	Ser	Gly	Ser	Lys 165	Arg	Lys	Ser	Asp	Ser 170	Ser	
		GAT 757	TCC	CAG	GAA	GGA	AAT	TAC	AAG	TCA	. GAA	GTG	AAC	AGC	: AAA	CCI	C AGG	
65			Ser	Gln	Glu 175	Gly	Asn	Tyr	Lys	Ser 180	Glu	Val	Asn	Ser	Lys 185	Pro	Arg	

AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA

SEQUENCE LISTING

5	(1) GENERAL INFORMATION:
5	(i) APPLICANT: Gorski, David H. Walsh, Kenneth
10	(ii) TITLE OF INVENTION: Growth Arrest Homeobox Gene
	(iii) NUMBER OF SEQUENCES: 4
15 20	<pre>(iv) CORRESPONDENCE ADDRESS: (A) ADDRESSEE: Calfee, Halter, and Griswold (B) STREET: 800 Superior Avenue (C) CITY: Cleveland (D) STATE: Ohio (E) COUNTRY: U.S.A. (F) ZIP: 44114-2688</pre>
	(v) COMPUTER READABLE FORM:
254	(A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible
	(C) OPERATING SYSTEM: PC-DOS/MS-DOS
25	(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
	(vi) CURRENT APPLICATION DATA:
- 200	(A) APPLICATION NUMBER: (B) FILING DATE:
30	(C) CLASSIFICATION:
	(viii) ATTORNEY/AGENT INFORMATION:
100 Total 100 To	(A) NAME: Golrick, Mary E.
35	(B) REGISTRATION NUMBER: 34829 (C) REFERENCE/DOCKET NUMBER: 22311/00114
4.0	(ix) TELECOMMUNICATION INFORMATION: (A) TELEPHONE: (216) 622-8200 (B) TELEFAX: (216) 241-0816 (C) TELEX: 980499
	(2) INFORMATION FOR SEQ ID NO:1:
45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2244 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: both (D) TOPOLOGY: linear
50	(ii) MOLECULE TYPE: cDNA
4	(iii) HYPOTHETICAL: NO
55	(iv) ANTI-SENSE: NO
60	(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 1971108
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
65	GTCAAGTGTT TATACGTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT

22	~	4	4	,	\sim	\sim	-	7	А
	•		1 .	,	11	ŧ 1	•	•	4

	CTGCTCAAAC CCGCGCGCT TTTACATTAG GAGTGAGTGG GGGAGAGTCC TAGGATTCT
5	AGTGAAAAGT GACAGCGCTT GGTGGACTTT GGGACCTTCG TGAAGTCTTC TGCTTGGAAG 180
	CTGAGACTTG CATGCC ATG GAA CAC CCC CTC TTT GGC TGC CTG CGC AGC 229
10	Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser 1 5 10
	CCC CAC GCC ACA GCG CAA GGC TTG CAC CCC TTC TCG CAG TCT TCT CTG 277
15	Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu 15 20 25
	GCC CTC CAT GGA AGA TCT GAC CAC ATG TCC TAC CCC GAA CTC TCC ACA 325
20	Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr 30 35 40
TO THE PARTY OF TH	TCT TCC TCG TCT TGC ATA ATC GCG GGA TAC CCC AAT GAG GAG GGC ATG 373
25	Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met 45 50 55
	TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC CAC CAC CAC
3.0	Phe Ala Ser Gln His His Arg Gly His His His His His His His 60 65 70 75
	CAT CAC CAC CAG CAG CAG CAG CAC CAG GCT CTG CAA AGC AAC TGG
3.5 3.0	His His His Gln Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp 80 85 90
Here May	CAC CTC CCG CAG ATG TCC TCC CCG CCA AGC GCG GCC CGG CAC AGC CTT 517
40	His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu 95 100 105
#**;	TGC CTG CAG CCT GAT TCC GGA GGG CCC CCG GAG CTG GGG AGC AGC CCT
45	Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro 110 115 120
	CCG GTC CTG TGC TCC AAC TCT TCT AGC CTG GGC TCC AGC ACC CCG ACC 613
50·	Pro Val Leu Cys Ser Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr 125 130 135
,	GGA GCC GCG TGC GCA CCA AGG GAT TAT GGC CGT CAA GCG CTG TCA CCC 661
55 55	Gly Ala Ala Cys Ala Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro 140 145 150 155
	GCA GAA GTG GAG AAG AGA AGT GGC AGC AAA AGA AAA AGC GAC AGT TCA
60	Ala Glu Val Glu Lys Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser 160 165 170
	GAT TCC CAG GAA GGA AAT TAC AAG TCA GAA GTG AAC AGC AAA CCT AGG 757
65	Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg 175 180 185

AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA

	805 Arg Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala 190 195 200
5	GAG TTC GCC CAT CAT AAC TAT CTG ACC AGA CTG AGA AGA TAT GAG ATA
	853 Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile 205 210 215
10	GCG GTG AAC CTA GAC CTC ACT GAA AGA CAG GTG AAA GTG TGG TTC CAG
	Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln 220 235 .
15	AAC AGG AGA ATG AAG TGG AAG CGG GTC AAG GGG GGA CAA CAA GGA GCT 949
	Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala 240 245 250
20	GCA GCC CGA GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA ACA CTT CTT
	Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu 255 260 265
25	CCA TCA GAG CTG TCA GGA ATT GGT GCA GCC ACC CTC CAG CAG ACA GGG
	Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly 270 275 280
36	GAC TCA CTA GCA AAT GAC GAC AGT CGC GAT AGT GAC CAC AGC TCT GAG
	Asp Ser Leu Ala Asn Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu 285 290 295
35	CAC GCA CAC TTA TGATACATAC AGAGACCAGC TCCGTTCTCA GGAAAGCACC 1145 His Ala His Leu 300
35	His Ala His Leu
35 mg mg mg 40	His Ala His Leu 300 ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG
35 	His Ala His Leu 300 ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205 CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC
35 mg mg mg 40	His Ala His Leu 300 ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205 CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265 ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA
46 45	His Ala His Leu 300 ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205 CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265 ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA 1325 AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA
46 45	His Ala His Leu 300 ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205 CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265 ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA 1325 AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA 1385 AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATTT TGACTCAAGT TGTCTCCAGA
45	His Ala His Leu 300 ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205 CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265 ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA 1325 AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA 1385 AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATTT TGACTCAAGT TGTCTCCAGA 1445 GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC
45	HI45 HIS Ala HIS Leu 300 ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205 CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265 ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA 1325 AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA 1385 AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATTT TGACTCAAGT TGTCTCCAGA 1445 GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC TGGGCAGGTA TTTGCTTTTG CTTGCACTGA AACTTAAACT GCTATCAAGT TAACCCCATGA
45	His Ala His Leu ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205 CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265 ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA 1325 AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA 1385 AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATTT TGACTCAAGT TGTCTCCAGA 1445 GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC TGGGCAGGTA TTTGCTTTTG CTTGCACTGA AACTTAAACT GCTATCAAGT TAACCCATGA AATAGTTTAT CTTGAACAGC CACAGTGCCT GAAATCACCA AGTGGATATA AAATGAACTG

1	7	Λ	
	•	*	-

TTTATACATT TTTTTGTCAG TCAGAGACAT CAGTATGTGC TTACTTGTGT TCAAGTAGAG 1805

GAAATGCAGT AGAGTCTGAT AGGACATATT CTTGGTACCA CAGACAAAAC AAATCTTCTG

TTGCATTGAC TATCAACTGC TGCAGATACA TTAGAGAACA CACCTAGCCC CCCTCCAGCC 1925

TCCCTCTGTT ATCGCTCGAA GACATTAGCG TCATAGGCAA GTAGTTACCT TGCCAAATGA 1985

GTCTTGTGTG GCAGATGTCT GATTTTGTAT CTTTAAACTG TTAATGGTAT GTGTCTGCTT 2045

CAGTTAACAG GGAAAAAGAT TTCTTCCTCA TTGTTTATGA TACAAAACCC AAGTGCCAAA 2105

CAAAGCTAGT TCTTCAAGGG ATAGATGAGA AACTGAATGT CTGACAAGTA GACTCAGCGA 2165

AAATACATTA TTTTTCAGAG GCTGTGTATT CATGCAGTAC AAGTCCTTGT ATTTTGTAAA 2225

AAAAAAGTT AAATAAATG 2244

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 303 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser Pro His Ala Thr Ala 1 5 .10 15

Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu Ala Leu His Gly Arg 20 25 30

Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr Ser Ser Ser Cys 35 40 45

Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met Phe Ala Ser Gln His 50 55 60

Gln Gln His Gln Ala Leu Gln Ser Asn Trp His Leu Pro Gln Met 85 90 95

Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu Cys Leu Gln Pro Asp 100 105 110

Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro Pro Val Leu Cys Ser 115 120 125

Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr Gly Ala Ala Cys Ala 130 135 140

30 3 3

45

50.

55

60

65

5

10

15

20

22311	/00114
-------	--------

	Pro 145	Arg	Asp	Tyr	Gly	Arg 150	Gln	Ala	Leu	Ser	Pro 155	Ala	Glu	Val	Glu	Lys 160
5_	Arg	Ser	Gly	Ser	Lys 165	Arg	Lys	Ser	Asp	Ser 170	Ser	Asp	Ser	Gln	Glu 175	Gly
 	Asn	Tyr	Lys	Ser 180	Glu	Val	Asn	Ser	Lys 185	Pro	Arg	Arg	Glu	Arg 190	Thr	Ala
10	Phe	Thr	Lys 195	Glu	Gln	Ile	Arg	Glu 200	Leu	Glu	Ala	Glu	Phe 205	Ala	His	His
15	Asn	Tyr 210	Leu	Thr	Arg	Leu	Arg 215	Arg	Tyr	Glu	Ile	Ala 220	Val	Asn	Leu	Asp
15	Leu 225	Thr	Glu	Arg	Gln	Val 230	Lys	Val	Trp	Phe	Gln 235	Asn	Arg	Arg	Met	Lys 240
20	Trp	ГЛа	Arg	Val	Lys 245	Gly	Gly	Gln	Gln	Gly 250	Ala	Ala	Ala	Arg	Glu 255	Lys
The state of the s	Glu	Leu	Val	Asn 260	Val	Lys	Lys	Gly	Thr 265		Leu	Pro	Ser	Glu 270	Leu	Ser
20-7	Gly	Ile	Gly 275	Ala	Ala	Thr	Leu	Gln 280	Gln	Thr	Gly	Asp	Ser 285	Leu	Ala	Asn
© 31 30	Asp	Asp 290	Ser	Arg	Asp	Ser	Asp 295	His	Ser	Ser	Glu	His 300	Ala	His	Leu	
	(2)	INFO	ORMA	PION	FOR	SEQ	ID 1	10:3:								4,
The state of the s		(i)	(1 (1	QUENC A) LE B) T' C) S' C) T'	ength CPE : Crani	nucl	ll ba Leic ESS:	ase p acid	pairs İ	;			•			
LO		(ii)	MOI	LECUI	LE T	PE:	CDN	A.								
	•	(iii)	HYI	POTHE	etic <i>i</i>	AL: 1	10									
15		(iv)) AN	ri—si	ense :	NO.										
50·		(ix)	(2	ATURI A) NI B) LO	AME/I			.941		٠						
		(xi)) SEG	QUENC	CE DI	ESCR	EPTIC	ON: S	SEQ I	D NO	0:3:					
1	GTC:	TTCT										CAC	CCG	СТС	փփփ	GGC
55	53														Phe	
50	TGC	CTG	CGC	AGC	CCT	CAC	GCC	ACG	GCG	CAA	. GGC	TTG	CAC	_	TTC	TCC
		Leu	Arg 10	Ser	Pro	His	Ala	Thr 15	Ala	Gln	Gly	Leu	His 20	Pro	Phe	Ser
55	CAA 149	TCC	TCT	CTC	GCC	CTC	CAT	GGA	AGA	TCT	GAC	CAT	' ATG	TC	TAC	c ccc
		Ser 25	Ser	Leu	Ala	Leu	His 30	Gly	Arg	Ser	Asp	His 35	Met	Ser	Tyr	Pro

	GAG CTC T	CT ACT TO	T TCC TC	A TCT TGC	ATA ATC	GCG GGA TA	C CCC AAC
	Glu Leu S		r Ser Sei			Ala Gly Tyr	
5	40		45		50		55
e ^{rre}	GAA GAG G 245	AC ATG TI	T GCC AG	C CAG CAT	CAC AGG	GGG CAC CA	C CAC CAC
10	Glu Glu A	sp Met Ph 6	e Ala Sei O	Gln His	His Arg 65	Gly His His	His His 70
10	CAC CAC C	AC CAT CA	C CAC CA	T CAG CAG	CAG CAG	CAC CAG GC	T CTG CAA
	293			Gln Gln		His Gln Ala	
15		_		80		85	
	341					CCG AGT GC	
20	Thr Asn T	p His Le 0	u Pro Gli	Met Ser 95	Ser Pro	Pro Ser Ala 100	Ala Arg
	CAT AGC C	TC TGC CT	C CAG CC	C GAC TCT	GGA GGG	CCC CCA GAG	G TTG GGG
		eu Cys Le	u Gln Pro	Asp Ser	Gly Gly	Pro Pro Glu 115	Leu Gly
25	AGC AGC C	CG CCC GT	C CTG TG	C TCC AAC	TCT TCC	AGC TTG GGG	כ ייירר אפר
	437					Ser Leu Gly	
30	120 .		125		130	ner ned GiA	ser ser 135
	ACC CCG A	T GGG GC	C GCG TG	ciece cce	GGG GAC	TAC GGC CG	C CAG GCA
3 5	Thr Pro Ti	or Gly Ala	a Ala Cys O	Ala Pro	Gly Asp. 145	Tyr Gly Arg	Gln Ala 150
11-14 m 11-15 m	CTG TCA C	CT GCG GA	G GCG GA	G AAG CGA	AGC GGC	GGC AAG AGG	G AAA AGC
40	-	:0 Ala Gl: 155	u Ala Glu	Lys Arg	Ser Gly	Gly Lys Arg 165	Lys Ser
. 40	GAC AGC TO	CA GAC TC	C CAG GA	A GGA AAT	TAC AAG	TCA GAA GTO	C AAC AGC
		er Asp Se:	r Gln Glu	Gly Asn	Tyr Lys	Ser Glu Val	Asn Ser .
45			a acc ac		300 333	GAG CAA AT	
	629				•		
50°·	185	.g min er	190		THE LYS	Glu Gln Ile 195	Arg Glu
30	CTT GAA G	CA GAA TT	T GCC CA	r cat aat	TAT CTC	ACC AGA CT	G AGG CGA
55	Leu Glu Ai 200	a Glu Ph	Pala His 205	His Asn	Tyr Leu 210	Thr Arg Leu	Arg Arg 215
55	TAC GAG A	a GCA GT	G AAT CT	G GAT CTC	ACT GAA	AGA CAG GT	A AAA GTC
	Tyr Glu I	e Ala Va. 22	l Asn Leu O	Asp Leu	Thr Glu 225	Arg Gln Val	Lys Val 230
60	TGG TTC C	AA AAC AG	G CGG AT	G AAG TGG	AAG AGG	GTA AAG GG	
	773					Val Lys Gly	
65		235	A 11G(240	-Ne wrd	val Lys Gly 245	
100	CAA GGA G	CT GCG GC	T CGG GA	A AAG GAA	CTG GTG	AAT GTG AA	A AAG GGA

22311/00114	22:	11	10.0	114
-------------	-----	----	------	-----

39

	7	20.	J /	U-U-I	7.4					33	,						
		Glr	Gly	Ala 250	Ala	Ala	Arg	Glu	Lys 255	Glu	Leu	Val	Asn	Val 260	ГĀЗ	Lys	Gly
5		ACA 869	CTT	CTC	CCA	TCA	GAG	CTG	TCG	GGA	ATT	GGT	GCA	GCC	ACC	CTC	CAG
			Leu 265	Leu	Pro	Ser	Glu	Leu 270	Ser	Gly	Ile	Gly	Ala 275	Ala	Thr	Leu	Gln
10		CAA 917	ACA	GGG	GAC	TCT	ATA	GCA	AAT	GAA	GAC	AGT	CAC	GAC	AG1	GAC	CAC
10			Thr	Gly	Asp	Ser	Ile 285	Ala	Asn	Glu	Asp	Ser 290	His	Asp	Ser	Asp	His 295
15		AGC 941	TCA	GAG	CAC	GCC	CAC	CTC	TGA								
	•		Ser	Glu	His	Ala 300	His	Leu									
20		(2)	INF	ORMA:	rion	FOR	SEQ	ID 1	10:4:	:							
				(i)	(A)	ENCE) LEM) TYM) TOM	NGTH:	: 302 mino	ami aci	ino a		,					
			(ii) 1	MOLE	CULE	TYPE	: pr	otei	.n							
30			. (:	xi) :	SEQUE	ence	DESC	RIPI	CION:	SEÇ) ID	NO:4					
		Met 1	Glu	His	Pro	Leu 5	Phe	Gly	Сув	Leu	Arg 10	Ser	Pro	His	Ala	Thr 15	Ala
<u> </u>		Gln	Gly	Leu	His 20	Pro	Phe	Ser	Gln	Ser 25	Ser	Leu	Ala	Leu	His 30	Gly	Arg
House Mr. carelle		Ser	Asp	His 35	Met	Ser	Tyr	Pro	Glu 40	Leu	Ser	Thr	Ser	Ser 45	Ser	Ser	Cys
40		Ile	Ile 50	Ala	Gly	Tyr	Pro	Asn 55	Glu	Glu	Asp	Met	Phe 60	Ala	Ser	Gln	His
45		His 65	Arg	Gly	His	His	His 70	His	His	His	His	His 75	His	His	His	Gln	Gln 80
		Gln	Gln	His	Gln	Ala 85	Leu	Gln	Thr	Asn	Trp 90	His	Leu	Pro	Gln	Met 95	Ser
50 [°]		Ser	Pro	Pro	Ser 100	Ala	Ala	Arg	His	Ser 105	Leu	Сув	Leu	Gln	Pro 110	Asp	Ser
4			Gly	115	7	·			120					125			
55		Ser	Ser 130	Ser	Leu	Gly	Ser	Ser 135	Thr	Pro	Thr	Gly	Ala 140	Ala	СЛа	Ala	Pro
60		Gly 145	Asp	Tyr	Gly	Arg	Gln 150	Ala	Leu	Ser	Pro	Ala 155	Glu	Ala	Glu	Lys	Arg 160
			Gly			165					170					175	
65		Tyr	. TAs	Ser	Glu 180	Val	Asn	Ser	Lys	Pro 185	Arg	ГÄв	Glu	Arg	Thr 190	Ala	Phe

Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His Asn

			195					200					205			
5	Tyr	Leu 210	Thr	Arg	Leu	Arg	Arg 215	Tyr	Glu	Ile	Ala	Val 220	Asn	Leu	Asp	Leu
<i>p</i> *	Thr. 225	Glu	Arg	Gln	Val	Lys 230	Val	Trp	Phe	Gln	Asn 235	Arg	Arg	Met	Lys	Trp 240
ro	Lys	Arg	Val	Lys	Gly 245	Gly	Gln	Gln	Gly	Ala 250	Ala	Ala	Arg	Glu	Lys 255	Glu
	Leu v	Val	Asn	Val 260	Lys	Lys	Gly	Thr	Leu 265	Leu	Pro	Ser	Glu	Leu 270	Ser	Gly
L5	Île	Gly	Ala 275	Ala	Thr	Leu	Gln	Gln 280	Thr	Gly	Asp	Ser	Ile 285	Ala	Asn	Glu
	Asp	Ser 290	His	qaA	Ser	qaA	His 295	Ser	Ser	Glu	His	Ala 300	His	Leu		